

NITISHA JAIN

Postdoctoral Research Associate, King's College London

Email: nitishajain15@gmail.com | Phone: +447767576799

 [/in/nitisha-jain](https://in/nitisha-jain) |  nitishajain.github.io |  github.com/nitishajain

SKILLS

- **TECHNICAL SKILLS AND AREAS:** Python | Java | C/C++ | NLP | Machine Learning | Responsible AI | Generative AI | Deep Learning | Data and Knowledge Engineering | Semantic Web | Explainable AI with neuro-symbolic methods | Knowledge Graphs | Natural Language Processing
- **FUNCTIONAL SKILLS:** Scientific paper writing | Project management | Research methodology | Cross-functional and cross-cultural collaboration | Analytical problem-solving | Strong verbal and written communication skills | Industry-academic partnerships | Leadership and Team work | Teaching/Mentoring PhD and Masters' students.

EXPERIENCE

[FEB 2023 – NOW]

POSTDOCTORAL RESEARCH ASSOCIATE, KING'S COLLEGE LONDON, UNITED KINGDOM

- *Mentors:* Prof. Dr. Elena Simperl, Dr. Albert Merono Penuela (Informatics Department)
- Developed pipeline for **multimodal content generation** (images, text, speech, 3D, music) from text using generative AI models (Stable Diffusion, MusicGen), combined with a **RAG pipeline over Wikidata** to guide generation; enabled scalable creation of 3,600+ multimodal representations for 500+ cultural items in the [MuseIT](#) EU Horizon project (2022–2025).
- Fine-tuned GPT-4 and gpt-3.5-turbo to automate knowledge graph completion and enrichment for Wikidata; combined **in-context learning, prompt engineering, and RAG techniques** to achieve a macro-averaged **F1-score of 0.701**, enhancing triple prediction accuracy and **winning LM-KBC challenge at ISWC 2023**.
- Collaborated with **leading industry researchers** (Google, Meta, HuggingFace, Kaggle) in MLCommons [Croissant](#) project to create a metadata standard for documenting Machine Learning datasets; led the design for the **Responsible AI vocabulary** for ensuring data interoperability and compliance with ethical AI guidelines for AI security and reliability; presented as **Spotlight paper at NeurIPS 2024**.
- Developed **neuro-symbolic methods** to improve interpretability in **graph embeddings**; created a framework aligning vectors with human-understandable semantic features leading to **improved transparency in semantic similarity tasks**, published at NeSy 2024 and **Neurosymbolic AI Journal 2025**.

[OCT 2020 – JAN 2021]

PHD INTERNSHIP, BOSCH CENTRE FOR ARTIFICIAL INTELLIGENCE (BCAI), GERMANY

- *Mentors:* Dr. Daria Stepanova, Dr. Trung Kien Tran (NLP and Semantic Reasoning Group)
- Applied **advanced algorithmic and reasoning techniques** to machine learning models for data representations, designed novel algorithm and architecture that improved semantic consistency and performance by more than **10%**.
- Conducted **cutting-edge research** resulting in first author publication at the **premier conference ISWC 2021**, demonstrating aptitude for research communication and scientific discovery.

- Enhanced collaboration skills through joint research projects, focusing on **practical implementations of AI research in industrial settings**.

[OCT 2017 – SEP 2018]

RESEARCH ASSISTANT, MAX PLANCK INSTITUTE FOR INFORMATICS, GERMANY

- Mentors: Dr. Simon Razniewski, Prof. Gerhard Weikum (Database & Information Systems Group)
- Employed **deep learning models** for measurable and recall-aware information extraction, contributing to the development of efficient knowledge engineering pipelines.
- Led the research project resulting in a paper at **EMNLP 2019** (rank A conference), showcasing innovative research for extracting valuable insights from large datasets on the web.
- Gained experience in **high-profile collaborative research environments**, working under the guidance of **leading experts** in the field.

[AUG 2015 – AUG 2017]

RESEARCH SCIENTIST, IBM RESEARCH AI, BANGALORE, INDIA

- Mentor: Arun Kumar, Senior Researcher and Manager (Cognitive Solutions)
- Developed **AI-driven solutions using IBM Watson technologies** for Cognitive Finance and Regulatory Compliance, enhancing automated decision-making processes.
- Constructed and optimized an information retrieval architecture for domain-specific documents.
- Built ontologies and knowledge graphs enabling more accurate and efficient data representation.
- Co-authored a demo paper at IUI 2017 and Technical Report in IT Professional, highlighting the practical **deployment of AI/ML solutions in an industrial context**.

[APR 2015 – MAY 2015]

RESEARCH INTERN, UNIVERSITY OF MELBOURNE, AUSTRALIA

- Mentor: Prof. Rajkumar Buyya (Clouds Lab, Department of CIS)
- Implemented deadline-based I/O scheduling on the CloudSim framework, improving resource allocation and efficiency in cloud computing environments.
- Facilitated **international research collaborations**, leading to first author **publication at CCBBD 2015**.

EDUCATION

[NOV 2018- OCT 2022]

PH.D. IN COMPUTER SCIENCE, HASSO PLATTNER INSTITUTE, UNIVERSITY OF POTSDAM, GERMANY

- Mentors: Prof. Dr. Ralf Krestel, Prof. Dr. Felix Naumann (Information Systems Group)
- Research Focus: Knowledge Graphs, Semantic Web, Information Retrieval, NLP, Machine Learning.
- **Thesis: Representation and Curation of Knowledge Graphs with Embeddings.**

[AUG 2012- MAR 2015]

MASTERS (RESEARCH) IN COMPUTER SCIENCE, INDIAN INSTITUTE OF SCIENCE, BENGALURU, INDIA

- Mentor: Dr. J. Lakshmi (Supercomputer Education and Research Centre)
- Research Focus: Operating Systems, Computer Architecture, Distributed Systems.
- Thesis: Performance Specific I/O Scheduling Framework for Cloud Storage.

[AUG 2007- JULY 2011]

BACHELORS IN INFORMATION TECHNOLOGY, GGSIP UNIVERSITY, NEW DELHI, INDIA

- Focus: Operating Systems, Computer Networks, Distributed Systems.

SELECTED PUBLICATIONS

- *Towards Interpretable Embeddings: Aligning Representations with Semantic Aspects. **Neurosymbolic AI Journal 2025**.
- *Croissant: A Metadata Format for ML-Ready Datasets. **Spotlight paper at NeurIPS 2024 (also Best Paper at DEEM, SIGMOD 2024)**.
- Towards deployment-centric multimodal AI beyond vision and language. **CoRR abs/2504.03603, 2025**.
- Talking Wikidata: Communication Patterns and Their Impact on Community Engagement in Collaborative Knowledge Graphs. **Transactions on Graph Data & Knowledge, 2025**.
- Agreeing and disagreeing in collaborative knowledge graph construction: An analysis of Wikidata. **Journal of Web Semantics, 2025**.
- A Standardized Machine-readable Dataset Documentation Format for Responsible AI. **CoRR abs/2407.16883, 2024**.
- Bringing back Semantics to Knowledge Graph Embeddings: An Interpretability Approach. **NeSy 2024**.
- *The Polifonia Ontology Network: Building a Semantic Backbone for Musical Heritage. **ISWC 2023**.
- Using Large Language Models for Knowledge Engineering (LLMKE): A Case Study on Wikidata. **Winner of KBC-LM Challenge, ISWC 2023**.
- *Discovering Fine-Grained Semantics in Knowledge Graph Relations. **CIKM 2022**.
- Generating Domain-Specific Knowledge Graphs - Challenges with Open Information Extraction. **Text2KG at ESWC 2022**.
- *Generation of Training Data for Named Entity Recognition of Artworks. **Semantic Web Journal 2021**.
- *Improving Knowledge Graph Embeddings with Ontological Reasoning. **ISWC 2021**.
- *Do Embeddings Actually Capture Knowledge Graph Semantics? **ESWC 2021**.
- *Coverage of Information Extraction from Sentences and Paragraphs. **EMNLP 2019**.

(*Rank A/A*)

SERVICE

- **Workshop Co-Organizer:** Generative Neuro-Symbolic AI (GeNeSy), ESWC 2024
- **Program Committee member:** Sponsorship Co-Chair, ESWC 2024
- **Reviewer:** ISWC 2025, Semantic 2025, Semantic Web Journal (2023–2025), SemDH 2025, Semantics 2024, TKDD 2023, The Web Conference 2022, ESWC 2022, CIKM 2022
- **Teaching & Supervision:**
 - TA: Knowledge Engineering (2025), Network Data Analysis (2024–2025)
 - Supervised intern (Jan–May 2024) on Explainable graph embeddings (resulted in 2 publications)